Statistical methods for marginal inference from multivariate ordinal data

Nooraee, Nazanin

DOI:
10.1016/j.csda.2014.03.009

Chapter 4

Patterns of Dupuytren Disease in Fingers: studying correlations with a multivariate ordinal logit model

R. Lanting, N. Nooraee, P.M.N. Werker, and E.R. van den Heuvel

Plastic and Reconstructive Surgery. 2014; 134(3): 483-489.

4.1 Abstract

Background: Dupuytren disease affects fingers in a variable fashion. Knowledge about specific disease patterns (phenotype) based on location and severity of the disease is lacking.

Methods: In this cross-sectional study, 344 primary affected hands with Dupuytren disease were physically examined. The Pearson correlation coefficient between the co-existence of Dupuytren disease in pairs of fingers was calculated, and agglomerative hierarchical clustering was applied to identify possible clusters of affected fingers. With a multivariate ordinal logit model we studied the correlation on severity, taking into account age and sex, and tested hypotheses on independence between groups of fingers.

Results: The ring finger was most frequently affected by Dupuytren disease, and contractures were seen in 15.1 percent of affected rays. The severity of thumb and index finger, middle and ring finger, and the middle and little finger was significantly correlated. Occurrences in pairs of fingers were highest in the middle and ring finger, and lowest in the thumb and index finger. Correlation between the ring and little finger, and a correlation between fingers from the ulnar and radial side could not be demonstrated.

Conclusions: Rays on the ulnar side of the hand are predominantly affected. The middle finger is substantially correlated with other fingers on the ulnar side, the thumb and index finger are correlated, however there was no evidence that the ulnar side and the radial side were correlated in any way, which suggests that occurrence on one side of the hand does not predict Dupuytren disease on the other side of the hand.

4.2 Introduction

Dupuytren disease is an incurable fibromatosis of the hand and fingers, giving rise to the development of skin pitting and subcutaneous nodules in the palm. At a later stage of disease cords appear that connect the nodules and may contract the fingers into a flexed position. A contracture can occur isolated in a single joint, but may also involve more joints of a single ray or even multiple rays, whereby the metacarpophalangeal joints (MCPJ), proximal interphalangeal joints (PIPJ), and distal interphalangeal joints (DIPJ) are affected in decreasing order. The disease is usually located on the ulnar side of the hand, and in particular the ring finger and the little finger are frequently affected[25,13,19].

Several authors have described the patterns of occurrence of Dupuytren disease in multiple fingers empirically. Meyerding noticed that the combination of an affected ring and little finger occurred most often, followed by the combination of an affected third, fourth, and fifth digit[18]. In addition, Tubiana has stated that isolated radial side involvement in Dupuytren disease is rare, and that radial involvement in most cases is associated with an affected ulnar side[29]. Milner et al. found that patients with a severely affected thumb which had required surgery, were on average eight years older, and had suffered significantly longer from Dupuytren disease than patients with a mildly affected radial side. Furthermore, these patients with severe disease of the thumb suffered from ulnar disease which repeatedly had required surgery, suggesting an intractable form of disease[19,1].

Orlando et al. reported that in most hands two rays are affected by Dupuytren disease, followed by one finger and three fingers. Affection of four or even five fingers was rarely seen[21]. A previous study of ours showed an average of 2.7 affected rays per patient[13].

In summary, the frequency in which each ray is affected has been described previously, as well as intra operative findings in relatively small samples[16], but firm statistical substantiation is lacking. The consequence is that certain findings may be spurious and not tested appropriately. For instance, the higher frequency of an affected ring and little finger may well be determined by the higher occurrences in these fingers, and may not be more frequent than expected by chance. This may be true also for other pairs, and additionally for triplets, quadruples, or even for the affection of all rays of one hand.

Furthermore, the existence of specific disease patterns (phenotype) based on location and severity of the disease has never been studied before. Therefore, the aim of this study is firstly to investigate the patterns of occurrence and severity of primary Dupuytren disease in both men and women, and secondly to test the earlier suggested correlations in Dupuytren disease occurrence between the little and ring finger, but more importantly between the ulnar and radial side of the hand.

4.3 Patients and Methods

4.3.1 Participants

To obtain a representative cross-sectional set of patients with Dupuytren disease, we included 105 patients from the general population of the northern Netherlands[13], as well as 134 patients from the outpatient clinic of the Department of Plastic Surgery of the University Medical Center Groningen. This study was approved by the institutional ethics review board.

4.3.2 Physical examination

Both hands of all included patients were physically examined by the first author (R.L.) at the outpatient clinic. Signs of Dupuytren disease, including presence of subcutaneous nodules and fascial cords, with or without finger contractures, were noted for each ray of each hand. The severity of flexion contractures was measured with a goniometer, and the passive extension deficit was noted in degrees for each joint separately. Thereafter, the severity of disease was categorized using the classification of Tubiana et al.[27], in which stage N refers to affection with only a nodus or cord without contracture[28] (Table 4.1).

4.3.3 Statistical analyses

Descriptive statistics of our population were calculated first. Furthermore, Pearson's correlation coefficient between the coexistence of Dupuytren disease among pairs of fingers in each hand separately was calculated and tested for its statistical significance. We performed a post-hoc power analysis, to calculate what minimal correlation we could detect with our sample size[3].

Table 4.1 Tubiana Classification.

Stage	Total Passive Extension Deficit (degrees)
0	No lesion
N*	0
1	1 – 45
2	46 – 90
3	91 – 135
4	>135

*Stage N refers to affection with only a nodus or cord without contracture.

To identify possible occurrence of patterns in fingers with Dupuytren disease, a hierarchical cluster analysis was conducted, assuming that patterns would be similar in both hands. The measure of similarity between fingers was based on Jaccard[4], and the complete-linkage method[26] was applied to form clusters of fingers. Agglomerative hierarchical clustering (from bottom to top) was used[8].

To investigate the influence of sex and age on the coexistence of Dupuytren disease, and to evaluate the patterns of severity, a multivariate ordinal logit model was fitted to the Tubiana stage of all five fingers simultaneously (assuming that patterns are similar in both hands again). For this statistical analysis we collapsed the three most severe categories of Tubiana into one category. This multivariate model is similar to a probit model[14], but it was altered to be able to fit logits instead of probits, so that the effects of age and sex have interpretations similar to logistic regression. Instead of correlations on the occurrence, this multivariate model provides correlations on the severity between fingers, corrected for covariates. To obtain confidence intervals on the parameter estimates of the multivariate model, bootstrapping of 1000 samples was applied. Based on the multivariate model, the predicted probabilities of occurrence of Dupuytren disease in multiple fingers simultaneously are presented, and compared to the probabilities based on independence. If fingers are affected completely independent from each other, it can be expected that they co-occur with a frequency that equals the product of the occurrence rates of the individually affected fingers. Independency on severity between the radial side and ulnar side was tested with the likelihood ratio test. The statistical terminology is explained in more detail online in A..

4.4 Results

In this study, data of 152 (63.6 percent) males and 87 (36.4 percent) females were used. Mean age of patients was 65.4 years (SD 9.8), and 344 hands were affected. The ulnar side of the hand was predominantly affected; the most frequently affected ray was the ring finger, followed by the little finger and middle finger (Figure 4.1).

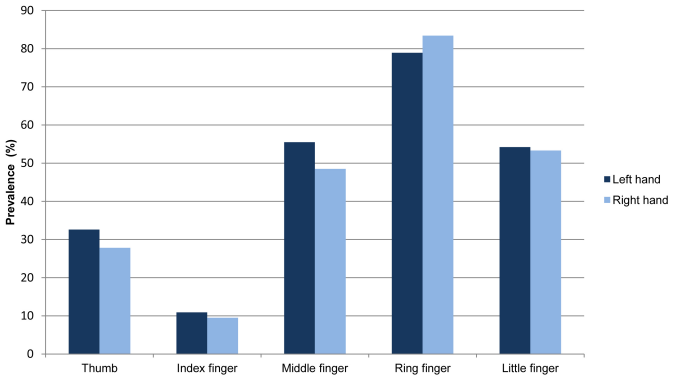

Fig. 4.1 Occurrence of rays affected with Dupuytren disease per hand.

The Pearson correlation coefficient for the coexistence of Dupuytren disease between fingers of both hands separately is provided in Table 2. We found that the thumb and index finger are significantly positively correlated in both hands (left 0.149; right 0.205), as well as the middle finger and the little finger of the left hand (0.262). Besides this, no high correlations were observed. Note that a correlation of 0.15 or higher could be detected with at least 80 percent power with the sample size of our cohort[3].

Assuming that the patterns would be the same across hands (which seems plausible considering the results from Figure 4.1 and Table 4.2), the dendrogram of the hierarchical clustering demonstrates that the middle finger and the ring finger should form the first cluster (Figure 4.2). This cluster was thereafter sequentially enlarged with the little finger, the thumb and the index finger. However, the short length(s) of the arms in the dendrogram suggest that the middle, ring, and little finger together form one cluster, while based on the longer length(s) of the arms, the thumb and the index finger should be seen as separate clusters. Thus three clusters were formed by the hierarchical clustering, where the ulnar side seems separate from the radial side.

Table 4.2 Pearson Correlation Coefficient for the Coexistence of Dupuytren Disease among Fingers on the Left Hand (lower triangle) and Right Hand (upper triangle).

	Thumb	Index Finger	Middle Finger	Ring Finger	Little Finger
Thumb	NA	0.205†	0.084	0.028	0.079
Index Finger	0.149*	NA	0.090	-0.019	0.141
Middle Finger	0.133	0.054	NA	0.019	0.127
Ring Finger	0.002	-0.134	0.099	NA	-0.035
Little Finger	0.026	0.025	0.262†	0.002	NA

NA, not applicable.
*Significant correlation at 0.05.
†Significa

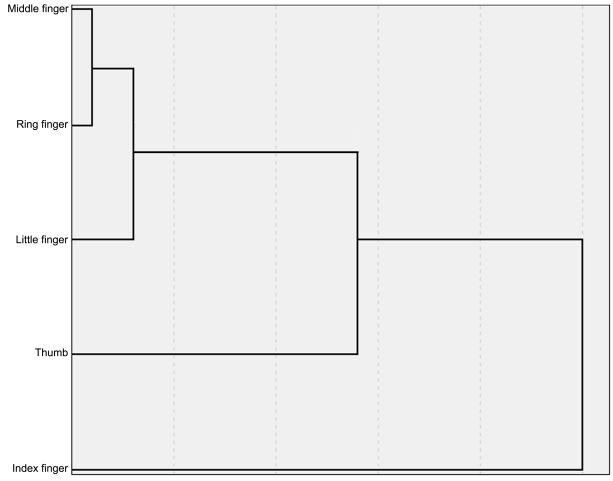

Fig. 4.2 Dendrogram of the hierarchical clustering of the occurrence of fingers using the Jaccard distance and complete linkage.

Table 4.3 shows the distribution of the severity of Dupuytren disease based on Tubiana stages. In most affected rays only nodules and cords were found, without an extension deficit. Contractures were seen in 15.1 percent of the affected rays.

Table 4.3 Frequency of Severity of Dupuytren Disease per Finger.

	Tubiana Stage			
	Unaffected (%)	N*	1 (%)	≥2 (%)
Thumb	240(69.8)	103(29.9)	1 (0.3)	-
Index Finger	309(89.8)	26(7.6)	8(2.3)	1(0.3)
Middle Finger	165(48.0)	159(46.2)	16(4.6)	4(1.2)
Ring Finger	65(18.9)	233(67.7)	44(12.8)	2(0.6)
Little Finger	159(46.2)	143(41.6)	32(9.3)	10(2.9)

*Stage N refers to affection with only a nodus or cord without contracture.

None of the identified patterns so far used information on the Tubiana stage,

and the results were not adjusted for covariates such as sex and age. Therefore, a multivariate logit model on severity—which takes age and sex into account— was fitted to the data. No effect of age (OR, 1.06; 95 [0.97 ; 1.19]) and sex (sex OR: 1.27 [0.98 ; 1.64]) on the severity of Dupuytren disease could be demonstrated. The correlation coefficients with 95 percent confidence intervals for severity of Dupuytren disease between pairs of fingers are presented in Table 4.4. There was a significant correlation between thumb and index finger, the middle and ring finger, and the middle and little finger. The significant positive correlations imply that those pairs occur more frequently than can be expected based on independence, and that a more severe disease of, for example, the middle finger is accompanied by more severe disease of the ring finger. The radial side (thumb and index) and the ring finger, as well as the ring finger and little finger, seem to be negatively correlated, although not significant.

Table 4.4 Correlation Coefficients with 95 Percent Confidence Intervals for Severity of Dupuytren Disease between Pairs of Fingers Based on the Multivariate Logit Model.

	Index Finger	Middle Finger	Ring Finger	Little Finger
Thumb	0.22†(0.01; 0.42)	0.16 (-0.04; 0.34)	-0.009(-0.24; 0.21)	0.09 (-0.08; 0.25)
Index Finger	NA	0.12 (-0.12; 0.35)	-0.17 (-0.39; 0.06)	0.14 (-0.06; 0.35)
Middle Finger		NA	0.29*(0.10; 0.47)	0.17†(0.02; 0.31)
Ring Finger			NA	-0.12 (-0.30; 0.05)

NA, not applicable.
*Significant correlation at 0.01.
†Significant correlation at 0.05.

The multivariate logit model makes it also possible to predict the occurrences of Dupuytren disease in single, pairs, triplets, quadruples, and quintets of rays for both genders at different ages. Since age only had a minor effect on the severity of Dupuytren disease in fingers, we report only the results at age 65 years (the average age of the sample). In Table 4.5, the highest predicted occurrence of Dupuytren disease in pairs of fingers was observed in the middle and ring finger at 46.2 percent (95 percent CI, [39.7 ; 53.0]) for men and 39.7 percent (95 percent CI, [32.8 ; 46.7] for women. Dupuytren disease in any pair of fingers seem to occur least frequent in the thumb and index finger at 5.4 percent (95 percent CI, [3.0 ; 8.3]) for men and 3.7 percent (95 percent CI, [1.9 ; 6.3]) for women. The highest predicted occurrence of all triplet combinations is a combination of the middle, ring and little finger with a occurrence of 29.5 percent (95 percent CI, [23.4 ; 35.4] for men and 23.2 percent (95 percent CI, [17.3 ; 29.5]) for women. For quintets the occurrence decreases to 2.1 percent (95 percent CI, [1.0 ; 3.5]) for men and 1.2

percent (95 percent CI, [0.4 ; 2.3]) for women. As expected from the occurrences of individual fingers and pairs, the triplet and quadruple combinations which include fingers from the ulnar side are more prevalent.

Table 4.5 Predicted Occurrences of Dupuytren Disease with 95% Confidence Intervals in Single and Combinations of Fingers for Men and Women Separately at Age 65 Years

	Occurrence, % (95% CI)	
Fingers*	Men	Women
One or two fingers		
1	30.9(25.6; 36.4)	26.1(20.8;31.6)
2	12.8(8.8; 17.4)	10.4(6.8;14.9)
3	53.8(47.8; 60.1)	48.7(41.3;54.3)
4	77.3(73.0; 81.1)	72.9(67.5;77.8)
5	59.5(53.0; 65.1)	53.6(46.7;59.9)
1, 2	5.4(3.0; 8.3)	3.7(1.9; 6.3)
1, 3	19.4(14.8; 24.1)	14.9(10.5;19.9)
1, 4	24.5(19.4; 29.6)	19.4(14.0;25.1)
1, 5	20.2(15.6; 25.1)	15.6(11.2;20.8)
2, 3	7.4(4.3; 11.0)	5.3(2.8; 8.3)
2, 4	8.3(5.1; 12.0)	5.8(3.1; 9.4)
2, 5	8.2(5.0; 12.0)	5.9(3.1; 9.3)
3, 4	46.2(39.7; 53.0)	39.7(32.8;46.7)
3, 5	35.5(29.4; 41.5)	29.2(23.1;35.5)
4, 5	45.9(39.5; 51.8)	38.7(31.5;45.8)
Three or more fingers		
1, 2, 3	3.7(1.9; 5.8)	2.4(1.1; 4.0)
1, 2, 4	3.9(2.0; 6.2)	2.4(1.0; 4.5)
1, 2, 5	3.9(2.1; 6.2)	2.5(1.2; 4.5)
1, 3, 4	16.3(11.9; 21.0)	12.1(7.9; 16.7)
1, 3, 5	13.3(9.4; 17.7)	9.5(6.1; 14.0)
1, 4, 5	15.5(11.5; 20.0)	11.2(7.2; 16.1)
2, 3, 4	5.7(3.1; 8.9)	3.8(1.8; 6.4)
2, 3, 5	5.4(2.9; 8.5)	3.6(1.7; 6.2)
2, 4, 5	5.6(3.2; 8.5)	3.7(1.7; 6.5)
3, 4, 5	29.5(23.4; 35.4)	23.2(17.3; 29.5)
1, 2, 3, 4	2.8(1.4; 4.7)	1.7(0.7; 3.1)
1, 2, 3, 5	2.8(1.3; 4.6)	1.7(0.7; 3.1)
1, 2, 4, 5	2.7(1.4; 4.5)	1.6(0.6; 3.2)
1, 3, 4, 5	11.0(7.6; 15.2)	7.5(4.5; 11.7)
2, 3, 4, 5	4.1(2.1; 6.7)	2.6(1.0; 4.6)
1, 2, 3, 4, 5	2.1(1.0; 3.5)	1.2(0.4; 2.3)

* 1, thumb; 2, index; 3, middle, 4, ring; 5, little.

The correlations in occurrence (Table 4.2) and severity (Table 4.4) in Dupuytren disease between the little finger and ring finger were not strong. We tested the hypothesis of independence by comparing the predicted occurrence of Dupuytren

disease in these fingers with the product of the predicted occurrences of the single fingers. The p-value was determined at p=0.491 for males and p=0.376 for females, which suggests that we cannot demonstrate a correlation between the little and ring finger. We also tested, in a similar fashion, whether the ulnar and radial side are independent. The predicted occurrence of the quintet is quite close to the product of the predicted occurrence of the triplet "little, ring, and middle finger" and the pair "thumb and index finger" (men, p=0.202; women, p=0.218). However, since the predicted occurrences are quite small, we also tested this hypothesis with the likelihood ratio test, by comparing our multivariate logit model with a similar multivariate logit model where all of the correlations on severity between fingers on the radial side and fingers on the ulnar side were set equal to zero. Again, independence between the ulnar side and radial side could not be rejected (LRT=9.53; df=6; p=0.146).

4.5 Discussion

The aim of this study was to scrutinize the phenotype (i.e. disease patterns) of primary Dupuytren disease in men and women. When studying the phenotype of Dupuytren disease, it is important to realize that supposed patterns are dependent on the number of times the individual rays are affected. For example, the ring finger and little finger are most frequently affected. As a consequence, these fingers will often be seen affected with Dupuytren disease in combination with other affected rays, and one could therefore inadvertently conclude that there is a disease pattern. However, true patterns only exist when they appear more often than is expected based on the individual frequencies. In our study, we have investigated this, in particular on patterns that were recognized in the past.

Several aspects should be noted regarding the etiology of the phenotype. The anatomy of the hand is very complex, and has been the subject of numerous publications. In addition, several authors have tried to elucidate which anatomical structures are affected in Dupuytren disease, Tubiana has written about Dupuytren disease on the radial side of the hand, where a close relation exists between the thumb and index finger through the distal and proximal transverse commissural ligaments[29]. Based on this anatomical finding it is conceivable that thumb and index finger affection is correlated. This correlation could be proven both in the occurrence, and with respect to the severity of Dupuytren disease.

Tubiana also suggested that isolated Dupuytren disease of the radial side is rare, however, we could not demonstrate that the radial and ulnar side were correlated. Thus, we conclude that the in past literature frequently observed occurrence of an affected radial side in combination with a single ray or multiple rays from the ulnar side, is just explained by the high occurrences of Dupuytren disease in fingers in the ulnar side, and low frequency in the radial side. Indeed, our results show that the ring finger ray is frequently affected, while the disease in the radial rays of the hand, especially the index finger, is relatively rare. Furthermore, in our sample with only primary affected hands, the minority of rays had a passive extension deficit. These findings on occurrence are in agreement with the results of previous research[19,10,2]. In our analyses we assumed that severity of disease does not affect the correlations between fingers. This assumption implies that late stage disease would show similar disease patterns as early stage disease, but with higher levels of severity. This is difficult to test, but the goodness-of-fit test of our statistical model did not show a lack-of-fit, which we interpreted as evidence for our assumptions. Therefore, we expect that a higher prevalence of contractures in our sample would not have changed the conclusions.

McGrouther has described different layers of longitudinal fibers in the distal palm, from which the most superficial fibers insert into the dermis of the distal palm and proximal phalanx. These fibers are especially prominent in the middle and ring finger[17], and this might be an explanation for the correlation between these two fingers. Besides, these fibers are also seen in the little finger, and this supports the correlation between the middle and little finger, and the high predicted occurrences of the triplet combination of the middle, ring and little finger.

Besides anatomical variations, at present unknown molecular abnormalities of the extracellular matrix (ECM) as well as cytokines and growth factors that are associated with pathogenesis of Dupuytren disease, may play a role in the phenotype. Although results of different studies are not unambiguous, it has been suggested that ECM-proteins, such as collagen, periostin and β-catenin stimulate the proliferation, differentiation, and invasiveness of fibroblasts[22,24]. Furthermore, oxidative stress is thought to be involved in the pathogenesis of Dupuytren disease[24]. Hypoxia activates the xanthine oxidase pathway, eventually resulting in formation of oxygen free radicals, which are thought to stimulate myofibroblast proliferation[22,20]. Based on these pathogenic processes, it is possible that areas in the hand that contain more ECM-proteins or are exposed to hypoxia, will be more

affected with Dupuytren disease. In this respect, environmental risk factors that are thought to be associated with Dupuytren disease, such as heavy manual work and exposure to vibration[5], may also affect the phenotype. It is conceivable that DD will be more present in areas in the hand where the largest forces are applied to. Since the little and ring finger are the fingers that are predominantly used to grip and hold objects, this might explain their more frequent affection, although we could not demonstrate a correlation between this pair of fingers. However, more fundamental research is needed to elucidate the effect of these external forces on various tissues. Furthermore, attention should be paid to the effect of the amount of force, the duration of exposure, and the recovery time between exposure.

Dolmans et al. identified genetic risk factors that contribute to Dupuytren disease. The presence of nine specific SNPs can be used to calculate a genetic risk score for each patient[7,6]. Patients with certain clinical findings, such as a positive family history or ectopic lesions, had a higher genetic risk score[6]. It would be interesting to study whether this genetic risk score is associated with certain disease patterns. For instance, whether patients that carry more risk alleles have a more extensive phenotype.

One of the strengths of this cross-sectional study is the large number of primary affected hands that were included. Furthermore, the severity of disease in all cases was measured by only one observer and categorized using the well known classification of Tubiana[27]. However, our sample represents patients from the northern Netherlands alone, and the findings of our study may not necessarily be transferable to other parts of the world. Indeed, it has been suggested that race and geographical location might play a role in the onset of the disease, and will therefore influence the prevalence of Dupuytren disease. On the other hand, there is no clear influence of race and geographical location on prevalence[12], which makes it even more difficult to understand the potential influence of geography on disease patterns and correlations. Previous publications from Europe[19,29] and Japan[1], showed a comparable distribution of occurrence of Dupuytren disease among fingers. Thus, we expect that our results will also be applicable to patients with Dupuytren disease from other countries, but this needs to be confirmed with other studies.

Previous studies on the phenotype of Dupuytren disease were most often observational studies without firm statistical analyses[19,18,29]. Therefore, our study with agglomerative hierarchical clustering and a multivariate ordinal logit model

adds great value to the existing literature.

In articles studying disease patterns or clusters of disease, hierarchical clustering is a frequently used method[15,23,9]. However, it is an explorative analysis, and there are several disadvantages to this method. Firstly, errors in clustering methods for binary data may be quite substantial, although an evaluation revealed that our method, using Jaccard's distance and complete linkage, is one of the best methods to choose[26]. Secondly, this analysis will forcedly create s clusters, even when in the data no natural clusters exist[11]. A third drawback is the inability to address risk factors with this explorative analysis.

In our multivariate model these disadvantages do not apply, and we included age and gender as covariates. Therefore, our results are applicable to a broad population. Although there was no statistically significant relation between severity of Dupuytren disease and gender and age, the results suggest that males and older patients have more severe disease. It has been stated that younger patients have a more aggressive form of the disease with a higher recurrence rate[30], so consequently these patients will have a more severe disease already at younger ages, reducing the overall effect of age on severity. It might have been interesting to include age of onset or duration of disease as well. However, we noticed that patients have difficulty remembering the exact age of onset of the disease, which makes use of this information unreliable.

4.6 Conclusions

Our study substantiates that the ulnar side of the hand is predominantly affected in Dupuytren disease. In addition, regarding severity of Dupuytren disease, we found a significant correlation between the thumb and the index finger, the middle and ring finger, and the middle and little finger. The ulnar and radial side of the hand do not seem to be significantly correlated in any way, nor could we demonstrate a correlation between the little and ring finger. Knowledge of these phenotypes is a first step towards further analysis of the role of the genotype in causing the various forms of Dupuytren.

Acknowledgment

This research was funded by the University Medical Center Groningen.

A. Appendix

Cluster analysis: Cluster analysis is a statistical technique to discover homogeneous subgroups based on a set of measurements. In practice, a cluster analysis is the end product of a series of analytical decisions. This series of analytic decisions typically involve choices about what objects to cluster, what proximity measure to use to determine similarity or dissimilarity among the objects, and what type of clustering algorithm to use[4]. In our data, we aimed to cluster fingers that were affected with Dupuytren disease. The measures of proximity and type of clustering algorithms are explained below.

Measures of proximity: To identify clusters of observations (i.e. combinations of affected fingers) in the data, it is important to know how similar individual observations (i.e. individual affected fingers) are, or how far apart they are. For binary data (a finger is affected with Dupuytren disease or not) several measures of similarity can be used, all based on measures of a 2×2 contingency table. We used Jaccard's coefficient. This method only gives weight to the similarity of two fingers when Dupuytren disease is present in these fingers. Fingers without Dupuytren disease are ignored in this similarity measure. Besides the proximity between two individual observations (for example affected thumb and affected little finger), it is important to measure the dissimilarity of groups of observations in a cluster analysis. This inter-group proximity is based on the inter-individual proximity. In the complete linkage cluster method, which we used, the inter-group dissimilarity is defined as the largest distance between two individual observations, one from each group[26]. This is also known as furthest-neighbor distance[8].

Type of clustering algorithm: Clustering algorithms can be classified as hierarchical or non-hierarchical. Hierarchical algorithms are most appropriate for classification when objects are related via some underlying systematic structure[4]. Hierarchical algorithms are further classified according to whether the algorithm proceeds by successively merging individual observations into groups (agglomerative method) or starts with one large cluster and separates the observations into smaller groups (divisive method). Agglomerative methods are most widely used[4,8].

In **agglomerative hierarchical clustering** the clusters are formed in several steps. It starts with all single objects as separate clusters, in our case all five fingers are seen as separate cluster at the beginning of the analysis. Successively these objects are grouped into larger clusters until the final grouping contains all the

original objects in one group. A dendogram illustrates which fusions are made in each step of the analysis[11]. For example, our dendogram (Figure 4.2 in the paper) shows that in the first step the middle finger and ring finger are clustered (and thus are most similar compared to other combinations). In the following step the little finger was added to this first cluster, keeping the thumb and index finger still as single clusters.

With a **multivariate ordinal logit model** an ordinal logistic regression-like analysis was performed. This model is suitable for categorical data with ordered categories (i.e. Tubiana stage), measured at multiple time points or locations (i.e. five fingers)[14]. The model takes into account that observations on one hand could be correlated. In addition, covariates can be included in the statistical model to distinguish in severity level between subgroups of patients. In our paper, we used this model to calculate the correlations on the severity between fingers, corrected for age and gender.

Bootstrapping: method to obtain confidence intervals on the parameter estimates of the multivariate model.

References

[1] Abe, Y., Rokkaku, T., Ofuchi, S., Tokunaga, S., Takahashi, K., and Moriya, H. (2004). Dupuytren's disease on the radial aspect of the hand: Report on 135 hands in japanese patients. *The Journal of Hand Surgery*, 29:359–362.

[2] Brenner, P., Krause-Bergmann, A., and Van, V. (2001). Dupuytren contracture in northern germany: Epidemiological study of 500 cases (in german). *Unfallchirurg*, 104(4):303–311.

[3] Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Elbraum Associates.

[4] Cornell, J., Pugh, J., Williams, J., Kazis, J., Lee, A., Parchman, M., Zeber, J., Pederson, T., Montgomery, K., and Noël, P. (2008). Multimorbidity clusters: Clustering binary data from multimorbidity clusters. clustering binary data from a large administrative medical database. *Applied Multivariate Research*, 12(3):163–182.

[5] Descatha, A., Jauffret, P., Chastang, J., Roquelaure, Y., and Leclerc, A. (2011). Should we consider dupuytren's contracture as workrelated? a review and meta-analysis of an old debate. *BMC Musculoskeletal Disorders*, 12(96).

[6] Dolmans, G., de Bock, G., and Werker, P. (2012). Dupuytren diathesis and genetic risk. *The Journal of Hand Surgery*, 37(10):2106–2111.

[7] Dolmans, G., Werker, P., Hennies, H., and et al. (2011). Wnt signaling and dupuytren's disease. *The New England Journal of Medicine*, 365:307–317.

[8] Everitt, B. (2011). *Cluster Analysis*. John Wiley & Sons.

[9] Formiga, F., Ferrer, A., Sanz, H., and et al. (2014). Patterns of comorbidity and multimorbidity in the oldest old: The octabaix study. *European Journal of Internal Medicine*, 24(1):40–44.

[10] Gudmundsson, K., Arngrìmsson, R., Sigfùsson, N., Björnsson, A., and Jónsson, T. (2000). Epidemiology of dupuytren's disease: Clinical, serological, and social assessment. the reykjavik study. *Journal of Clinical Epidemiology*, 53(3):291–296.

[11] Jain, A. and Dubes, R. (1988). *Algorithms for Clustering Data*, chapter Clustering methods and algorithms, page 55–142. Prentice Hall.

[12] Lanting, R., Broekstra, D., Werker, P., and van den Heuvel E.R. (2014). A systematic review and meta-analysis on the prevalence of dupuytren disease in the general population of western countries. *Plastic and Reconstructive Surgery*, 133(3):593–603.

[13] Lanting, R., van den Heuvel, E., Westerink, B., and Werker, P. (2013). Prevalence of dupuytren disease in the netherlands. *Plastic and Reconstructive Surgery*, 132(2):394–403.

[14] Li, Y. and Schafer, D. (2008). Likelihood analysis of the multivariate ordinal probit regression model for repeated ordinal responses. *Computational Statistics & Data Analysis*, 52(7):3474–3492.

[15] Marengoni, A., Rizzuto, D., Wang, H., Winblad, B., and Fratiglioni, L. (2009). Patterns of chronic multimorbidity in the elderly population. *Journal of the American Geriatrics Society*, 57(2):225–230.

[16] McFarlane, R. (1974). Patterns of the diseased fascia in the fingers in dupuytren's contracture: Displacement of the neurovascular bundle. *Plastic and Reconstructive Surgery*, 54(1):31–44.

[17] McGrouther, D. *Dupuytren's Disease Biology and Treatment*, chapter The palm, pages 127–135. Churchill Livingstone.

[18] Meyerding, H. (1936). Dupuytren's contracture. *Archives of Surgery*, 32:320–333.

[19] Milner, R. (2003). Dupuytren's disease affecting the thumb and first web of the hand. *The Journal of Hand Surgery*, 28(1):33–46.

[20] Murrell, G.A. Francis, M. and Bromley, L. (1990). Modulation of fibroblast proliferation by oxygen free radicals. *Biochemical Journal*, 265(3):659–665.

[21] Orlando, J., Smith, J., and Goulian, D. (1974). Dupuytren's contracture: A review of 100 patients. *British Journal of Plastic Surgery*, 27:211–217.

[22] Picardo, N. and Khan, W. (2012). Advances in the understanding of the aetiology of dupuytren's disease. *Surgeon*, 10(3):151–158.

[23] Sakagami, T., Hasegawa, T., Koya, T., and et al. (2014). Cluster analysis identifies characteristic phenotypes of asthma with accelerated lung function decline. *Journal of Asthma*, 51(2):113–118.

[24] Shih, B. and Bayat, A. (2010). Generalized estimating equations for correlated binary data: Using the odds ratio as a measure of associationscientific understanding and clinical management of dupuytren disease. *Nature Reviews Rheumatology*, 6(12):715–726.

[25] Stahl, S. and Calif, E. (2008). Dupuytren's palmar contracture in women. *The Israel Medical Association journal*, 10(6):445–447.

[26] Tamasauskas, D., Sakalauskas, V., and Kriksciuniene, D. (2012). Evaluation framework of hierarchical clustering methods for binary data. *aper presented at: 12th International Conference on Hybrid Intelligent Systems (HIS); December 4–7; Pune, India:*, page 421–426.

[27] Tubiana, R., Michon, J., and Thomine, J. (1968). Scheme for the assessment of deformities in dupuytren's disease. *Surgical Clinics of North America*, 48(5):979–984.

[28] Tubiana, R., Michon, J., and Thomine, J. (1974). *Evaluation of deformity in Dupuytren's Contracture*. Churchill Livingstone.

[29] Tubiana, R., Simmons, B., and DeFrenne, H. (1982). Location of dupuytren's disease on the radial aspect of the hand. *Clinical Orthopaedics and Related Research*, 168:222–229.

[30] van Rijssen A.L., ter Linden, H., and Werker, P. (2012). Five-year results of a randomized clinical trial on treatment in dupuytren's disease: Percutaneous needle fasciotomy versus limited fasciectomy. *The American Statistician*, 129(2):469–477.